Lifetime Data Analysis, 8, 147–161, 2002
© 2002 Kluwer Academic Publishers. Printed in The Netherlands.

Testing Equality of Cause-Specific Hazard Rates Corresponding to *m* Competing Risks Among *K* Groups

S. B. KULATHINAL
Department of Epidemiology and Health Promotion, National Public Health Institute, Mannerheimintie 166, 00300 Helsinki, Finland

DARIO GASBARRA
Rolf Nevanlinna Institute, University of Helsinki, P.L.4 00014 Helsinki, Finland

Received March 14, 2000; Revised June 13, 2001; Accepted June 28, 2001

Abstract. In this paper, a class of tests is developed for comparing the cause-specific hazard rates of m competing risks simultaneously in K (≥ 2) groups. The data available for a unit are the failure time of the unit along with the identifier of the risk claiming the failure. In practice, the failure time data are generally right censored. The tests are based on the difference between the weighted averages of the cause-specific hazard rates corresponding to each risk. No assumption regarding the dependence of the competing risks is made. It is shown that the proposed test statistic has asymptotically chi-squared distribution. The proposed test is shown to be optimal for a specific type of local alternatives. The choice of weight function is also discussed. A simulation study is carried out using multivariate Gumbel distribution to compare the optimal weight function with a proposed weight function which is to be used in practice. Also, the proposed test is applied to real data on the termination of an intrauterine device.

Keywords: cause-specific hazard, competing risks, counting processes, cumulative incidence function, martingale limit theorem

1. Introduction

Consider the situation where m competing risks are acting simultaneously in the same environment. For each unit subject to m competing risks, the data available are the failure time of the unit, U and the identifier of the risk claiming the failure, ϵ which can assume any one of the possible values $\{1, 2, \ldots, m\}$. In practice, U is censored and data on $T = min(U, C)$ are available, where C is a censoring variable and the censoring mechanism is assumed to be noninformative. Also, instead of ϵ, $\delta = \epsilon I[T = U]$ is observed. We assume that the data available are (T, δ), right censored competing risks data. Suppose there are K independent groups of units, and each unit is exposed to m competing risks. Let T_{ki} be the failure time and $\delta_{ki} \in \{0, 1, 2, \ldots, m\}$ be the cause of failure, $i = 1, 2, \ldots, n_k$ and $k = 1, 2, \ldots, K$. For each k, the pairs (T_{ki}, δ_{ki}), $i = 1, 2, \ldots, n_k$, are assumed to be independent and identically distributed. It is important to note that no assumption regarding the dependence among the m risks is made.

Define the cumulative incidence function for the risk j in the group k by

$$F_{kj}(t) = P[U_{ki} \le t, \epsilon_{ki} = j], \tag{1}$$

which are assumed to be continuous with subdensities $f_{kj}(t)$. Also define the cause-specific hazard rate by

$$h_{kj}(t) = \frac{f_{kj}(t)}{S_k(t)},$$

where $S_k(t) = P[U_{kj} > t]$ is an overall survival function of the group k. Note that $S_k(t) = 1 - F_k(t) = 1 - \sum_{j=1}^{m} F_{kj}(t)$.

It is important to note that the cause-specific hazard rates for risk j, $j = 1, 2, \ldots, m$ corresponding to (U, ϵ) are the same as those corresponding to (T, δ).

The main purpose of this paper is to develop a test procedure for the hypothesis

$$H_o : h_{1j}(t) = h_{2j}(t) = \ldots = h_{Kj}(t) = h_j(t), \ \forall j = 1, 2, \ldots, m,$$

where $h_j(.)$, $j = 1, 2, \ldots, m$ are unspecified common cause-specific hazard rates.

Analysis of competing risks data concentrates on estimating and comparing cause-specific hazards, as can be seen from previous works (see Kalbfleisch and Prentice, 1980). Hence it is of direct interest to propose a test procedure which makes use of cause-specific hazards.

The problem of testing equality of cause-specific hazard rates corresponding to m dependent risks has been discussed in the literature (see Lam, 1998; Aly et al., 1994 and references therein). Gray (1988) gives a class of $K-$ sample tests for comparing the cumulative incidence functions of a particular type of failure out of several competing risks among different groups and is based on the subdistribution hazards $(f_{kj}(t)/(1 - F_{kj}(t)))$. Lindkvist and Belyaev (1998) propose a class of tests based on a two-dimesional vector statistic for testing equality of cumulative cause-specific hazard rates corresponding to two risks between two samples. Their approach can be easily applied to m risks.

In section 2, a class of K-sample tests, a generalization of Lindkvist and Belyaev (1998) test, is developed using martingale theory. The asymptotic distribution of the test statistic, which is univariate unlike in Lindkvist and Belyaev (1998), is shown to be a chi-squared distribution. The choice of the weight function in the light of maximizing asymptotic efficiency against a sequence of local alternatives is discussed in section 3. The technical details are given in the appendix.

In section 4, the proposed test is applied to simulated data from multivariate Gumbel distribution to compare the optimal weight function with the proposed weight function which is to be used in practice. Also, the empirical distribution under the null hypothesis is plotted and the empirical power is computed against the local alternatives.

The motivation for this work came from the analysis of contraceptive failure data in intrauterine device (IUD) studies carried out in several countries. The possible causes of termination of the use of IUD were categorised into five risk classes: (1) pregnancy, (2) expulsion, (3) amenorrhea, (4) bleeding and pain, and (5) hormonal disturbances. In the present context, it is essential and interesting to study simultaneously the behaviour of all five of the above-mentioned competing termination reasons simultaneously for different countries. None of the available tests can be directly applied to this situation.

2. Test Statistic and Asymptotic Distribution

Suppose (T_{ki}, δ_{ki}), $i = 1, 2, \ldots, n_k$, $k = 1, 2, \ldots, K$ denote right censored competing risks data for the K groups. Let $n = \sum_{k=1}^{K} n_k$. Define the counting process

$$N_{kj}(t) = \sum_{i=1}^{n_k} I[T_{ki} \leq t, \delta_{ki} = j], \; j = 0, 1, 2, \ldots, m; k = 1, 2, \ldots, K.$$

and

$$Y_k(t) = \sum_{i=1}^{n_k} I[T_{ki} \geq t].$$

Note that $N_{kj}(t)$ counts the number of failures due to competing risk j by time t and $Y_k(t)$ is the number of units at risk just prior to time t for the k^{th} group. Set $M_{kj}(t) = N_{kj}(t) - \int_0^t Y_k(s)\, d\Lambda_{kj}(s), j = 0, 1, 2, \ldots, m, k = 1, 2, \ldots, K$. Then for $t \in [0, \tau]$, $M_{kj}(t), j = 0, 1, 2, \ldots, m$, $k = 1, 2, \ldots, K$ are orthogonal square integrable martingales with respect to the filtration $\{\mathcal{F}_t^{N,Y}\}$ which is generated by N_{kj} and Y_k.

Define the overall counting process

$$N_{\cdot j}(t) = \sum_{k=1}^{K} N_{kj}(t)$$

$$Y_{\cdot}(t) = \sum_{k=1}^{K} Y_k(t),$$

and the martingale

$$M_{\cdot j}(t) = \sum_{k=1}^{K} M_{kj}(t) = N_{\cdot j}(t) - \sum_{k=1}^{K} \int_0^t Y_k(s) d\Lambda_{kj}(s),$$

for each $j = 0, 1, \ldots, m$.

Note that

$$\langle M_{kj}, M_{k'j'} \rangle = \delta_{kk'}\delta_{jj'} \int_0^t Y_k(s) d\Lambda_{kj}(s), \text{ and } \langle M._j, M_{kj'} \rangle = \delta_{jj'} \langle M_{kj}, M_{kj'} \rangle.$$

A K-sample test can be based on the scores for $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, K$

$$Z_{ki} = \int_0^{\tau_k} \sum_{j=1}^{m} K_{kij}^n(t) \Big\{ d\hat{\Lambda}_{kj}(t) - d\hat{\Lambda}_j(t) \Big\}, \tag{2}$$

where $K_{kij}^n\ (t)$ are suitably chosen locally bounded $\{\mathcal{F}_t^{N,Y}\}$-predictable processes. $\hat{\Lambda}_{kj}(t)$, an estimator of the cumulative cause-specific hazard rate for competing risk j in the group k, is given by

$$d\hat{\Lambda}_{kj}(t) = \frac{dN_{kj}(t)}{Y_k(t)} = \frac{dM_{kj}(t)}{Y_k(t)} + d\Lambda_{kj}(t),$$

and $\hat{\Lambda}_j(t)$ is an estimate of the common value of $\Lambda_{kj}(t)$ under the null hypothesis, given by

$$d\hat{\Lambda}_j(t) = \frac{dN._j(t)}{Y.(t)} = \frac{dM._j(t)}{Y.(t)} + \frac{\sum_{k=1}^{K} Y_k(t) d\Lambda_{kj}(t)}{Y(t)}.$$

When $K_{kij}^n(t) = \delta_{ij} K_{ki}^n(t)$ then (2) simplifies to

$$Z_{ki} = \int_0^{\tau_k} K_{ki}^n(t) \Big\{ d\hat{\Lambda}_{ki}(t) - d\hat{\Lambda}_i(t) \Big\}, \tag{3}$$

which is shown to be a generalisation of the test proposed by Lindkvist and Belyaev (1998). When $m = 2$ and $K = 2$, the score Z_{11} can be simplified as

$$\begin{aligned} Z_{11} &= \int_o^{\tau_k} K_{11}^n(t) \left\{ \frac{dN_{11}(t)}{Y_1(t)} - \frac{dN_{11}(t) + dN_{21}(t)}{Y_1(t) + Y_2(t)} \right\} \\ &= \int_0^{\tau_k} K_{11}^n(t) \frac{Y_2(t)}{Y_1(t) + Y_2(t)} \left\{ \frac{dN_{11}(t)}{Y_1(t)} - \frac{dN_{21}(t)}{Y_2(t)} \right\}. \end{aligned}$$

A similar expression can be given for Z_{12}. The expression (3) on page 145 in the paper by Lindkvist and Belyaev (1998) is similar to the above expression with $V_{jn}(t) = K_{1j}^n(t)$ $\frac{Y_2(t)}{Y_1(t)+Y_2(t)}$ which has the form of the optimal weight function given on page 147 with $K_{1j}^n(t) = c_j^{-1} Y_1(t)$.

The martingale central limit theorem can be applied in this context to derive the asymptotic distribution of Z (Andersen et al., 1993).

THEOREM 2.1 *Assume $n^{-1}Y_k^n(s) \to y_k(s)$ uniformly in probability, where $y_k(s)$ are deterministic functions. Denote $y.(s) = \sum_{k=1}^{K} y_k(s)$. Then, under the assumption that $n^{-1}K_{kij}^n(t) \to K_{kij}^o(t)$ uniformly in probability, with each $K_{kij}^o(.)$ bounded on $[0, \tau]$ and under the null hypothesis, as $n \to \infty$, $n^{-1/2}Z$ converges in distribution to $N_{m\times K}(\underline{0}, \sum)$, where*

$$\Sigma_{(i,k),(i',k')} = \int_0^{\min(\tau_k,\tau_{k'})} \left(\frac{\delta_{kk'}}{y_k(t-)} - \frac{1}{y.(t-)} \right) \sum_{j=1}^{m} K_{kij}^o(t) K_{k'i'j}^o(t) d\Lambda_j(t). \tag{4}$$

Proof: Expressing Z_{ki}^n in terms of martingales, we have

$$\begin{aligned} Z_{ki}^n &= \int_0^{\tau_k} \sum_{j=1}^{m} K_{kij}^n(t) \left[\frac{dN_{kj}(t)}{Y_k^n(t)} - \frac{dN._j(t)}{Y.^n(t)} \right] \\ &= \int_0^{\tau_k} \sum_{j=1}^{m} K_{kij}^n(t) \left[\frac{dM_{kj}(t)}{Y_k^n(t)} - \frac{dM._j(t)}{Y.^n(t)} \right] \\ &\quad + \int_0^{\tau_k} \sum_{j=1}^{m} K_{kij}^n(t) \left[d\Lambda_{kj}(t) - \frac{\sum_{k=1}^{K} Y_k^n(t) d\Lambda_{kj}(t)}{Y.^n(t)} \right], \end{aligned}$$

where the second term is zero under the null hypothesis.

The result follows by standard arguments, since

$$\langle Z_{ki}^n, Z_{k'i'}^n \rangle = \int_0^{\min(\tau_k,\tau_{k'})} \left(\frac{\delta_{kk'}}{Y_k^n(t-)} - \frac{1}{Y.^n(t-)} \right) \sum_{j=1}^{m} K_{kij}^n(t) K_{k'i'j}^n(t) d\Lambda_j(t). \qquad \square$$

The consistent estimator of $\Sigma_{(i,k)(i'k')}$ is given by

$$n^{-1} \int_0^{\min(\tau_k,\tau_{k'})} \left(\frac{\delta_{kk'}}{Y_k^n(t-)Y.^n(t-)} - \frac{1}{Y.^n(t-)^2} \right) \sum_{j=1}^{m} K_{kij}^n(t) K_{k'i'j}^n(t) dN._j(t),$$

where $\tau = min(\tau_k, \tau_l)$.

Note that, when $K_{kij}^n(t) = \delta_{ij} K_{ki}^n(t)$, the asymptotic covariance matrix becomes block-diagonal, $\Sigma = \text{diagonal}(D_1, D_2, \ldots, D_m)$ where $D_i(k, k') = \Sigma_{(i,k),(i,k')}$, $i = 1, \ldots, m$, $k, k' = 1, \ldots, K$.

Under the alternative, as $n \to \infty$, $n^{-1/2} Z$ converges in distribution to $N_{m\times K}(\underline{\mu}, \Sigma)$, where

$$\mu_{ki} = \int_0^{\tau_k} \sum_{j=1}^{m} K^o_{kij}(t) \left[d\Lambda_{kj}(t) - \frac{\sum_{k=1}^{K} y_k(t) d\Lambda_{kj}(t)}{y.(t)} \right].$$

3. A Class of Tests

To generate a class of tests of H_o, we take the weight process of the form

$$K^n_{kij}(t) = L^n_{ij}(t) Y^n_k(t), i, j = 1, 2, \ldots, m, k = 1, 2, \ldots, K. \tag{5}$$

For the weight process of this type $\sum_{k=1}^{K} Z_{ki} = 0$ for each i. Hence, in this case, the rank of Σ is $(K - 1)$ rank(A), where A is the $m \times m$ matrix

$$A(i,j) = \int_0^{\tau} L^o_{ij}(t) h_j(t) dt,$$

where $n^{-1} K^n_{kij}(t)$ tends to $L^o_{ij}(t) y_k(t)$ as $n \to \infty$. In the follow-up we assume that A has full rank m. Hence, under the null hypothesis, the test statistic $n^{-1} Z'\Sigma^- Z$, where Σ^- is a generalised inverse of Σ, has asymptotically chi-squared distribution with $m(K-1)$ degrees of freedom. Under the alternative, the test statistic has asymptotically noncentral chi-squared distribution with $m(K-1)$ degrees of freedom with the noncentrality parameter $\underline{\mu}'\Sigma^-\underline{\mu}$. We show that the locally asymptotic efficient nonparametric test belongs to this class.

Consider a sequence of local asymptotic alternatives

$$h^n_{kj}(t) = h_j(t) + a_n^{-1} h_j(t) \sum_{i=j}^{m} \gamma_{ij}(t) \phi_{ki} + o(a_n^{-1}),$$

for $k = 1, 2, \ldots, K$ and $j = 1, 2, \ldots, m$, and $a_n = n^{1/2}$ throughout. The motivation for this alternative comes from the Gumbel's distribution which is illustrated in the next section.

We follow the technique given on pages 615-624 of Andersen et al. (1993) to show that the nonparametric test is asymptotically equivalent to an efficient parametric test. The proof is given in the appendix.

When the matrices $(\gamma_{ij}(t))_{t \geq 0}$ can be diagonalized by the same linear transformation simultaneously for all t, the sequence of local alternatives can be expressed as

$$h^n_{kj}(t) = h_j(t) + a_n^{-1} \phi_{kj} \gamma_j(t) h_j(t) + o(a_n^{-1}), j = 1, 2, \ldots, m; k = 1, 2, \ldots, K,$$

where ϕ_{kj} are constants and $\gamma_j(t)$ are fixed functions, $j = 1, 2, \ldots, m$, $k = 1, 2, \ldots, K$.

Note that $\gamma_j(t) \equiv 1$ gives the local alternative

$$h_{kj}(t) = h_j(t)(1 + a_n^{-1}\phi_{kj} + o(a_n^{-1})),$$

which corresponds to the proportionality of cause-specific hazards. The null hypothesis of equality corresponds to $\phi_{kj} = 0$ for all k and j. A test based on the process $L_j^n(t) = I[Y(t) > 0]$ corresponds to the log-rank test, in the absence of competing risks and is optimal when there is no censoring. An attractive process in our case is the type of process suggested by Harrington and Fleming (1982). We will consider the weight process

$$L_j^n(t) = [1 - \hat{F}_j(t)]^\rho, \tag{6}$$

where ρ is a fixed constant between 0 and 1 and $\hat{F}_j(t)$ is an estimate of the common incidence function for risk j, $P[U \leq t, \epsilon = j]$ and is given by

$$\hat{F}_j(t) = \int_0^t \hat{S}(u-)\frac{dN_j(u)}{Y(u)}, \tag{7}$$

where $\hat{S}(t-)$ is the left-hand limit of the Kaplan-Meier (1958) estimate of the survival function of U. When there is only one risk and censoring, $F_j(t) = F(t)$ and the above weight process is equivalent to the Harrington and Fleming (1982) type of tests for censored survival data. We refer to Chapter V and VIII of Andersen et al. (1993) for details regarding the tests for censored survival data. The corresponding $\gamma_j(t)$ process is $(1-F_j(t))^\rho$. Of course, in the present situation one can use the weight process $[\hat{S}(t)]^\rho$ which does not vary with j, which will give an optimal test for the class of alternatives with $\gamma_j(t) = S(t)^\rho$ for all j. Depending on the choice of $\gamma_j(.)$ process, one can get the optimal test by suitably defining $L_j^n(.)$.

In practice, the structure of the alternative is not known. In such situations it is easier and sufficient to use one of the weight functions given above. The illustration given in the following section brings out that the weight function in (6) performs satisfactorily when compared with the optimal weight function.

4. Simulation Study - Multivariate Gumbel Distribution

We consider a m-variate Gumbel exponential distribution, with parameters α and $\lambda = (\lambda_1, \ldots, \lambda_m)$, with the density

$$f(x_1, x_2, \ldots, x_m) = \prod_{i=1}^m \lambda_i \exp\left(-\sum_{i=1}^m \lambda_i x_i\right)\left[1 + \alpha \prod_{i=l}^m \left(2\exp(-\lambda_i x_i) - 1\right)\right].$$

The cause-specific hazard for the i-th risk is given by

$$h_i(t) = \lambda_i \frac{1 + \alpha(1 - \exp(-2\lambda_i t)) \prod_{j \neq i}(1 - \exp(-\lambda_j t))}{1 + \alpha \prod_{j=1}^{m}(1 - \exp(-\lambda_j t))}.$$

Now suppose we have a sequence of local alternatives $P^{(n)}$ with parameters $\alpha^{(n)} = \alpha$, and different parameters $(\lambda_{k1}^n, \ldots, \lambda_{km}^n)$ for each group, given by

$$\lambda_{ki}^n = \lambda_i + a_n^{-1} b_{ki} = \lambda_i(1 + a_n^{-1}\phi_{ki}),$$

where $a_n^{-1} \to 0$ and $\phi_{ki} = b_{ki}/\lambda_i$. Then

$$\begin{aligned}\frac{h_{ki}^n(t)}{h_i(t)} = 1 + a_n^{-1}\Bigg\{ \phi_{ki} + &\frac{\alpha \prod_{j=1}^{m}(1 - \exp(-\lambda_j t))}{1 + \alpha \prod_{j=1}^{m}(1 - \exp(-\lambda_j t))} \left(\sum_{j=1}^{m} \frac{\exp(-\lambda_j t)}{1 - \exp(-\lambda_j t)} \lambda_j \phi_{kj} t \right) \\ &-\alpha \prod_{j \neq i}(1 - \exp(-\lambda_j t)) \Big[1 + \alpha(1 - \exp(-2\lambda_i t)) \\ &\times \prod_{j \neq i}(1 - \exp(-\lambda_j t))\Big]^{-1} \exp(-2\lambda_i t) 2\lambda_i \phi_{ki} t + (1 - \exp(-2\lambda_i t)) \\ &\times \sum_{j \neq i} \Big[\exp(-\lambda_j t)\lambda_j \phi_{kj} t (1 - \exp(-\lambda_j t))^{-1}\Big] \Bigg\} \\ &+ o(a_n^{-1}) = 1 + a_n^{-1} \sum_{j=i}^{m} \gamma_{i,j}(t)\phi_{kj} + o(a_n^{-1}).\end{aligned}$$

We will consider $K = 5$ groups and $m = 3$ risks, and the true weight function corresponding to the optimal test and also the weight function given in (6) with $\rho = 0.5$. The level of significance used throughout is 0.05. The null hypothesis is rejected if the test statistic is greater than 21.026. The parameters used in the simulation are $\alpha = 0.5$ throughout and $\lambda = (0.8, 0.2, 0.6)$ for the null hypothesis. The censoring distribution is taken as exponential for each group with intensities (1, 0.6, 0.7, 0.8, 0.9), respectively. Samples of sizes (50, 60, 70, 80, 90) were generated for five groups with 1000 repetitions. Figure 1 gives the empirical distribution of the test statistic under the null hypothesis for the two weight functions along with the true chi-squared distribution with 12 degrees of freedom. The empirical distributions are quite close to the true distribution for moderate sample sizes. The empirical level of significance using the optimal weight function is 0.056 while it is 0.06 when the weight function (6) is used.

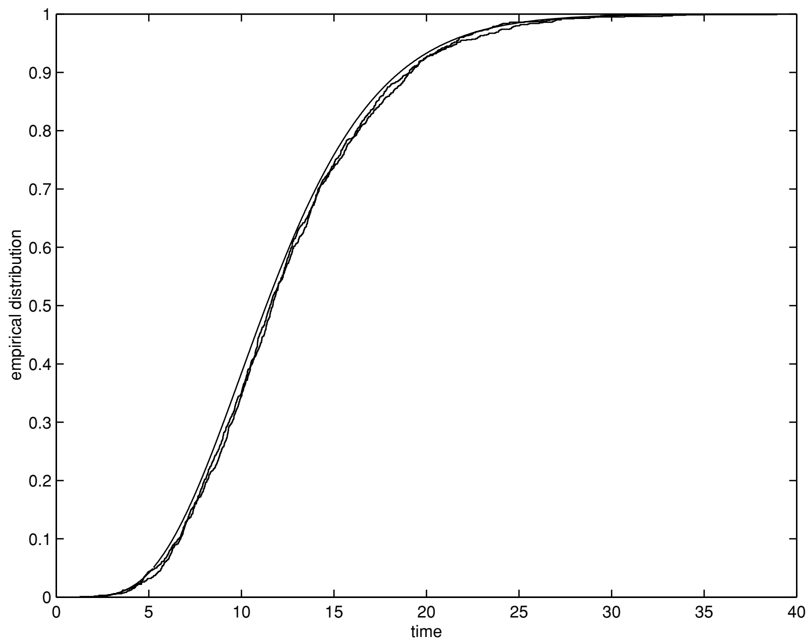

Figure 1. Empirical distributions of the asymptotically efficient test, the one based on the Harrington and Fleming type weight function (6) under the null hypothesis and the chi-squared distribution with 12 degrees of freedom (3 risks, 5 groups, multivariate Gumbel distribution).

To check the performance of the test as well as to compare the two weight functions, noncentrality parameters and empirical powers are computed when the parameters are (0.8, 0.2, 0.6), (0.79, 0.19, 0.59), (0.789, 0.21, 0.61), (0.75, 0.21, 0.55), and (0.81, 0.15, 0.52) for the five groups. Table 1 gives the noncentrality parameters and powers for various sample sizes. The empirical power of the test is 0.78 when optimal weight function is used and is 0.77 when Harrington and Fleming (1982) type of weight function is used for the sample sizes (200, 240, 280, 320, 360). The test performs satisfactorily even when the parameters are close to each other. It is clear that the Harrington and Fleming (1982) type of weight function can be used in practice since it gives the power which is reasonably close to the power of the test when the optimal weight function is used.

5. Application - IUD Study

The data are taken from a five year follow-up study of 1547 women from Finland, Sweden and Hungary, on termination of IUD conducted by a pharmaceutical company based in Finland. Here, termination of IUD due to (1) pregnancy, (2) expulsion, (3) amenorrhea, (4) bleeding and pain, and (5) hormonal disturbances was of interest. Table 2 gives the summary of the data used here.

Table 1. Noncentrality parameters and empirical powers of the test proposed in section 2 using optimal and Harrington and Fleming (HF) type weight functions for multivariate Gumbel distribution, level of significance = 0.05.

	noncentrality		power	
sample sizes	(optimal)	(HF)	(optimal)	(HF)
(50, 60, 70, 80, 90)	4.44	4.28	0.21	0.20
(100, 120, 140, 160, 180)	8.94	8.79	0.452	0.450
(200, 240, 280, 320, 360)	16.96	15.96	0.78	0.77

Table 2. Summary of IUD data.

Termination due to	Finland	Sweden	Hungary
pregnancy	2	0	0
expulsion	31	28	9
amenorrhea	11	7	50
bleeding and pain	60	96	64
hormonal disturbances	46	80	12
censoring	398	430	223
total	548	641	358

Figure 2 gives the estimates of the cause-specific hazard rates in the three countries for various termination reasons. The cause-specific hazards seem to vary between the countries. It can be seen from these plots that Finland and Sweden behave similarly with respect to these termination reasons except pregnancy but Hungary differs from these two Nordic countries. Sweden and Hungary reported no termination due to pregnancy and hence the cause-specific hazard rate is zero.

The main interest was in testing the equality of the five cause-specific hazards for the three countries. The weight functions used were the same as in (6) with $\rho = 0, 0.5, 1$. The test statistics for each risk are (3.27, 4.08, 106.28, 12.51, 23.21), (3.27, 4.11, 106.23, 12.69, 23.2) and (3.27, 4.13, 106.17, 12.86, 23.18) respectively.

The values of the test statistic are 149.38, 149.52, and 149.63 respectively. These values are higher than the cut-off point, 18.3, of the chi-squared distribution with 10 degrees of freedom at 0.05 level of significance. Hence, the hypothesis of equality of cause-specific hazards is rejected.

The cut-off point of chi-squared distribution with 2 degrees of freedom is 5.99. It can be seen that the first two reasons do not differ significatively among the countries, while the termination due to amenorrhea differs highly significantly between Hungary and the Nordic countries, Finland and Sweden. It was pointed out in Karia et al. (1998) that the termination due to amenorrhea depends on counselling and also on how often it is regarded as being disturbing by the user or her doctor. The opposite is true about the termination due to hormonal disturbances.

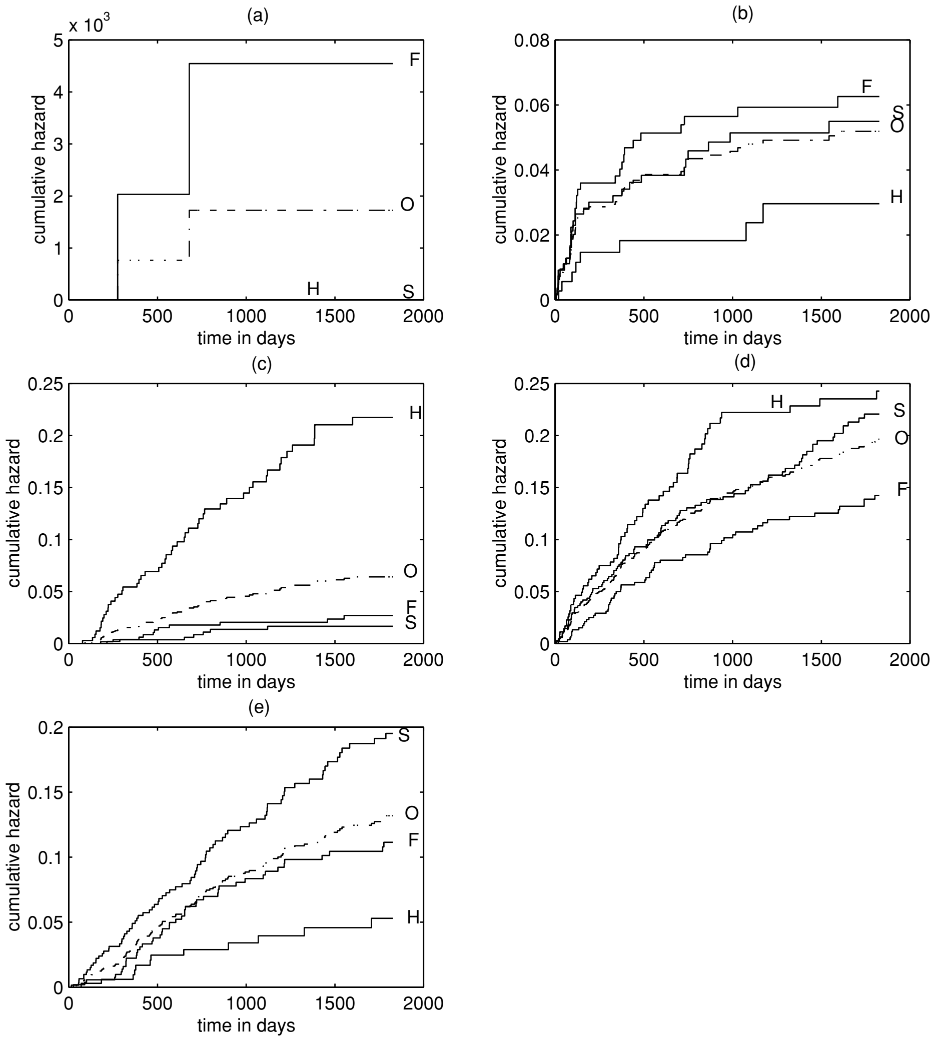

Figure 2. Estimates of cumulative cause-specific hazards for (a) pregnancy, (b) expulsion, (c) amenorrhea, (d) bleeding and pain, and (e) hormonal disturbances, in Finland (F), Sweden (S) and Hungary (H). Also the overall cumulative hazard (O) is given.

6. Remarks

A test for comparing cause-specific hazard rates proposed in this paper is very general in nature since almost all the available tests can be seen as special cases. The test seems to perform satisfactorily for moderate sample sizes. One can use the Harrington and Fleming type of weight function (6) in practice since its performance in a very general set-up is close to the performance of the optimal weight function.

Appendix

The local alternative corresponds to a parametric model

$$h^n_{kj}(t) = h_j(t, \underline{\theta}^n_k)$$

where

$$\underline{\theta}^n_k = (\theta^n_{k1}, \ldots, \theta^n_{km}) \in \mathbb{R}, \text{ with } \theta^n_{ki} = \theta^0_i + a_n^{-1}\phi_{ki}, \text{ and } h_j(t) = h_j(t, \underline{\theta}^0).$$

Note that

$$\gamma_{ij}(t) = \frac{\partial}{\partial\theta_i} \log h_j(t, \underline{\theta}^0).$$

The score for θ_{ki} at $\underline{\theta}_0$ is given by

$$U_{ki} = \sum_{j=1}^{m} \int_0^{\tau} \frac{\partial}{\partial\theta_i} \log h_j(t, \theta^0) dM_{kj}(t).$$

The vector $U = (U_{11}, \ldots, U_{1m}, \ldots, U_{K1}, \ldots, U_{Km})$ has asymptotically normal distribution with block diagonal $(mK) \times (mK)$ covariance matrix

$$\mathcal{I}_{(ki),(k'i')} = \delta_{kk'} \sum_{j=1}^{m} \int_0^{\tau} \gamma_{ij}(t)\gamma_{i'j}(t) y_k(t) h_j(t, \underline{\theta}^0) dt.$$

Note that if $(\gamma_{ij}(t))$ is diagonal, then $\mathcal{I}$ is diagonal.

Let $\phi \coloneqq (\phi_{11}, \ldots, \phi_{1m}, \ldots, \phi_{K1}, \ldots, \phi_{Km})' \in \mathbb{R}^{m\times K}$, and $\mathcal{J}$ is a $m \times m$ matrix with

$$(\mathcal{J})_{i,i'} = \sum_{j=1}^{m} \int_0^{\tau} \gamma_{ij}(t)\gamma_{i'j}(t) y.(t) h_j(t, \underline{\theta}^0) dt,$$

i.e., $\mathcal{J} = \sum_{h=1}^{K} \mathcal{I}_h$, where the $m \times m$ matrix $\mathcal{I}_h$ is the h-th block of $\mathcal{I}$. The noncentrality parameter for ϕ is given by

$$\zeta = \phi' \left\{ \mathcal{I} - \mathcal{I} \begin{pmatrix} \mathcal{J}^{-1} & \dots & \mathcal{J}^{-1} \\ \dots & \dots & \dots \\ \mathcal{J}^{-1} & \dots & \mathcal{J}^{-1} \end{pmatrix} \mathcal{I} \right\} \phi$$

$$= \sum_{l=1}^{K} (\phi^l)' \mathcal{I}_l \phi^l - \overline{\phi}' \left(\sum_{l=1}^{K} \mathcal{I}_l \right)^{-1} \overline{\phi}$$

where $\phi^l = (\phi_{l1}, \dots, \phi_{lm})'$, and $\overline{\phi} \in \mathbb{R}^{\mathrm{m}}$ is defined as

$$\overline{\phi} = \sum_{l=1}^{K} \mathcal{I}_l \phi^l$$

If the risk sets are asymptotically proportional, meaning that $y_l(t)/y.(t) = p_l$ constant, then

$$\zeta = \phi'((\mathrm{diag}(p) - p'p) \otimes \mathcal{J})\phi$$

$$= \sum_{i=1}^{m} \sum_{i'=1}^{m} \left[\left(\sum_{l=1}^{k} p_l \phi_{li} \phi_{li'} \right) - \overline{\phi}_i \overline{\phi}_{i'} \right] \sum_{j=1}^{m} \int_o^{\tau} \gamma_{ij}(t) \gamma_{i'j}(t) y.(t) h_j(t, \underline{\theta}^0) dt$$

where $p = (p_1, \dots, p_K)$ and

$$\overline{\phi}_i = \sum_{l=1}^{K} p_l \phi_{li}.$$

This is the optimal parametric test under the parametric model $\{h_i(t, \underline{\theta}), i = 1, \dots, m, \underline{\theta} \in \Theta\}$. It is easy to check that under the asymptotic proportionality of risk sets, the noncentrality parameters of the nonparametric test coincides with that of the efficient parameteric test and hence it is efficient. But the above parametric model does not allow for all possible shapes of the hazard rates and does not prove that the proposed nonparametric test is efficient in general. The above discussion helps to understand the derivation of noncentrality parameter. To show that the nonparametric test is efficient in general, we extend the above model and introduce a $K \times m$ dimensional nuisance parameter $\underline{\psi} = (\psi_{ki} :\ i = 1, 2, \dots, m, k = 1, 2, \dots, K)$. The local parameterisation of nuisance parameter is $\psi_{ki}^n = \psi_{ki}^0 + a_n^{-1} \eta_{ki}$ and we assume that

$$\frac{\partial}{\partial \psi_{ki}} \log h_j(t, \underline{\theta}^0, \underline{\psi}^0) = \gamma_{ij}(t) \frac{y_k(t)}{y.(t)}.$$

This corresponds to the parametric model

$$h_j(t, \underline{\theta}_k^n, \underline{\psi}^n) = h_j(t, \underline{\theta}_k^n) + h_j(t, \underline{\theta}_k^0) \sum_{i=1}^{m} \sum_{l=1}^{K} \gamma_{ij}(t) \frac{y_l(t)}{y.(t)} (\psi_{li}^n - \psi_{li}^0),$$

where $h_j(t, \underline{\theta}_k^n)$ is the original model and $\underline{\psi}$ varies in an open neighbourhood of $\underline{\psi}^0$. Note that the term comprising nuisance parameters is the same for all (k, j) and hence the null and alternative hypotheses remain the same.

The score for ψ_{ki} at $\underline{\psi}_0$ is given by

$$S_{ki} = \sum_{l=1}^{K} \sum_{j=1}^{m} \int_0^{\tau} \frac{\partial}{\partial \psi_{ki}} \log h_j(t, \underline{\theta}^0, \underline{\psi}^0) dM_{lj}(t)$$

and the effective information matrix is $(mK) \times (mK)$ dimensional given by

$$(\mathcal{I}^{\phi\phi|\psi})_{(li),(l'i')} = \int_0^{\tau} \sum_{j=1}^{m} \gamma_{ij}(t)\gamma_{i'j}(t) \frac{y_l(t)}{y.(t)} \left(\delta_{ll'} - \frac{y_{l'}(t)}{y.(t)} \right) y.(t) h_j(t, \underline{\theta}^0, \underline{\psi}^0) dt$$

and the noncentrality parameter is given by

$$\zeta = \phi' \mathcal{I}^{\phi\phi|\psi} \phi = \int_0^{\tau} \sum_{i}^{m} \sum_{i'}^{m} \left\{ \left(\sum_{l=1}^{K} \phi_{li}\phi_{li'} \frac{y_l(t)}{y.(t)} - \overline{\phi}_i(t)\overline{\phi}_{i'}(t) \right) \right.$$
$$\left. \times \sum_{j=1}^{m} \gamma_{ij}(t)\gamma_{i'j}(t) h_j(t, \underline{\theta}^0, \underline{\psi}^0) \right\} y.(t) dt$$

where $\overline{\phi}_i(t) := \sum_{l=1}^{K} \frac{y_l(t)}{y.(t)} \phi_{li}$.

We consider now the performance of the nonparametric chi-squared test $n^{-1} Z' \Sigma^{-} Z$ based on the Z_{ki}'s and Σ given in Theorem 2.1, where in the weight function (5) we have $L_{ij}^n(t) \to \gamma_{ij}(t)$ as $n \to \infty$, uniformly in probability.

It follows that

$$\Sigma = \mathcal{I}^{\phi\phi|\psi} \text{ and } \mu = \mathcal{I}^{\phi\phi|\psi} \phi. \tag{8}$$

Therefore, the noncentrality parameter of the nonparametric test is

$$\mu' \Sigma^{-} \mu = \phi' \mathcal{I}^{\phi\phi|\psi} (\mathcal{I}^{\phi\phi|\psi})^{-} \mathcal{I}^{\phi\phi|\psi} \phi$$
$$= \phi' \mathcal{I}^{\phi\phi|\psi} \phi.$$

Hence, the given nonparametric chi-squared test is locally asymptotically efficient.

Acknowledgments

We would like to thank the Associate editor and reviewers for their constructive comments and helpful suggestions. We acknowledge Leiras OY, which provided the data. A part of this paper was done while the first author was at the Department of Statistics, Panjab University, Chandigarh, India. The second author would like to acknowledge the support of the Academy of Finland.

References

E. A. A. Aly, S. C. Kochar, and I. W. McKeague, "Some tests for comparing cumulative incidence functions and cause-specific hazard rates," *J. Am. Statis. Assoc.* vol. 89 pp. 994–999, 1994.

P. K. Andersen, O. Borgan, R. D. Gill, and N. Keiding, *Statistical Models Based on Counting Processes*, Springer: New York, 1993.

R. L. Gray, "A class of K-sample tests for comparing the cumulative incidence of a competing risk," *Ann. Statist.* vol. 16 pp. 1141–1154, 1988.

D. P. Harrington and T. R. Fleming, "A class of rank test procedures for censored survival data," *Biometrika*, vol. 69 pp. 133–144, 1982.

J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, Wiley: New York, 1980.

E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *J. Am. Statis. Assoc.* vol. 53 pp. 457–481, 1958.

S. R. Karia, J. Toivonen, and E. Arjas, "Analysis of contraceptive failure data in intrauterine device studies," *Contraception* vol. 58 pp. 361–374, 1998.

K. F. Lam,"A class of tests for the equality of K cause-specific hazard rates in a competing risks model," *Biometrika* vol. 85 pp. 179–188, 1998.

H. Lindkvist and Y. Belyaev, "A class of non-parametric tests in the competing risks model for comparing two samples," *Scand. J. Statis.* vol. 25 pp. 143–150, 1998.